An interpretable axiomatization of the Hirsch-index

Denis Bouyssou[1] and Thierry Marchant[2]

[1] *bouyssou@lamsade.dauphine.fr*
Université Paris Dauphine - CNRS, Place du Maréchal de Lattre de Tassigny, F-75775 Paris Cedex 16 (France)

[2] *thierry.marchant@ugent.be*
Ghent University, H. Dunantlaan 1, 9000 Ghent (Belgium)

Introduction

In the last ten years, many bibliometric indices have been proposed for comparing and/or evaluating scientists. Among theses indices, the Hirsch-index (or h-index) is probably the most popular one. Since it is not clear which index is best, some researchers have tried to enrich the debate by analyzing various indices from an axiomatic perspective. This stream of research has delivered six[i] axiomatizations of the h-index: Woeginger (2008a,b), Quesada (2009), Quesada (2010), Quesada (2011), Miroiu (2013). They pave the way towards a better understanding of the h-index, but they are not completely satisfactory. That is why we propose a new axiomatization.

Existing axiomatizations and their shortcomings

Consider an index h' defined as 100 times the h-index. Is it worse or better than the h-index ? This question is obviously irrelevant, just like asking whether measuring distances in meter is better than in centimeters. Unfortunately, all aforementioned papers axiomatize the h-index instead of considering the family of all indices h' such that h' is equal to k times h. The axioms in these papers are therefore stronger then needed: they implicitely, or sometimes explicitely, state that the h-index of a scientist with one publication and one citation is one, while this actually does not matter.
We now discuss some specific problems.

Woeginger (2008a)

Theorem 4.1 in Woeginger (2008) characterizes the h-index by three axioms called A1, B and D. Axiom A1 is stated as follows: “If the $(n+1)$-dimensional vector y results from the n-dimensional vector x by adding a new article with $f(x)$ citations, then $f(y) \leq f(x)$.” Although this axiom is mathematically fine, we claim that it is not interpretable. Indeed, an axiom is a condition imposed on the index f, where f is any index, not necessarily the h-index. So, when we read this condition, we may not suppose that f is the h-index. It could be the square of the number of papers or the logarithm of the total number of citations, … It does therefore not make sense to say “if we add a new paper with $f(x)$ citations, then …” Why would we find such a condition (normatively) appealing if we do not know what $f(x)$ represents ? Axiom D has the same problem.

Woeginger (2008b)

This paper assumes that a bibliometric index must be a non-negative integer. This is very restrictive and difficult to motivate. It also uses axiom A1 as in Woeginger (2008a).

Quesada (2009)

Here, Axiom A1 imposes that $f(x)$ lies between (a) the minimum of the number of cited papers and the smallest number of citations (not taking uncited papers into account), and (b) the minimum of the number of papers and the largest number of papers. This is a complex condition. Actually, it combines several conditions.

Miroiu (2013)

This paper also assumes that a bibliometric index must be a non-negative integer. Besides, it uses some axioms (CPI, PR, CCI and CJ) that suffer the same problem as axiom A1 in Woeginger (2008a): they compare an unspecified index to a number of citations. This is not interpretable as long as we do not know which index is considered.

A new axiomatization

Among the aforementioned axiomatizations, those of Quesada seem most promising. We propose hereunder a list of axioms, inspired from those of Quesada, and we use them to axiomatize the family of all indices h' such that h' is equal to k times h.

Non-Triviality: there are scientists x, y such that $f(x) \neq f(y)$.

Zero: scientists with no paper or only uncited papers have an index equal to 0.

Tail Independence: suppose x and y have the same number of papers and $f(x) = f(y)$. Suppose both publish an additional paper, with the same number of citations, at most equal to the number of citations of the least cited paper of x and y. Then $f(x') = f(y')$.

Square Upwards: suppose x has m papers, each with m citations. Suppose x gets some additional citations. Then $f(x') = f(x)$.

Square Rightwards: suppose x has m papers, each with m citations. Suppose x publishes some additional papers with at most m citations. Then $f(x') = f(x)$.

Homothety: suppose x has m papers, each with m citations, and y has one paper, with one citation. Then $f(x) = m\, f(y)$.

Theorem : an index f satisfies Non-Triviality, Zero, A2 (Quesada), Tail Independence, Square Upwards, Square Rightwards and Homothety iff f is the h-index multiplied by some positive real number.

Compared to Theorem 3.1 in Woeginger (2009), our Theorem is more interesting because it axiomatizes the family of all h-indices. Moreover, it uses simpler axioms. For instance, A1 has been splitted into Square Upwards and Square Rightwards.

References

Burgos, A. (2010). Ranking scientists. *UMUFAE Economics Working Papers, Universidad de Murcia*, 2.

Gagolewski, M. (2011). Possibilistic analysis of arity-monotonic aggregation operators and its relation to bibliometric impact assessment of individuals. *International Journal of Approximate Reasoning*, 52, 1312-1324.

Marchant, T. (2009). An axiomatic characterization of the ranking based on the h-index and some other bibliometric rankings of authors. *Scientometrics*, 80, 325-342.

Miroiu, A. (2013). Axiomatizing the Hirsch index: Quantity and quality disjoined. *Journal of Informetrics*, 7, 10-15.

Quesada, A. (2009). Monotonicity and the Hirsch-index. *Journal of Informetrics*, 3, 158-160.

Quesada, A. (2010). Further characterizations of the Hirsch index. *Scientometrics*, 87, 107-114.

Quesada, A. (2011). More axiomatics for the Hirsch index. *Scientometrics*, 82, 413-418.

Woeginger, G. H. (2008a). An axiomatic characterization for the Hirsch-index. *Mathematical Social Sciences*, 56, 224-232.

Woeginger, G. H. (2008b). A symmetry axiom for scientific impact indices. *Mathematical Social Sciences*, 2, 298-303.

---

[i] Notice that Marchant (2009) does not belong to this list because it does not axiomatize the h-index, but the ranking induced by the h-index. Burgos (2010) and Gagolewski (2011) do also not belong to the list because they do not axiomatize the h-index but a family of indices containing the h-index.